Did U.S. Politicians Expect the China Shock?

Matilde Bombardini, Bingjing Li
and Francesco Trebbi*
[PRELIMINARY AND INCOMPLETE]

September 2020

Abstract

The China Shock has had a profound economic impact across regions of the US in the two decades straddling China's entry into the WTO. Two fundamental questions, which apply to arguably any policy decisions, are whether politicians knew about the extent of the China Shock's economic consequences in their district and how much those constituent interests affected their vote. We employ repeated roll call votes in the US Congress on China's Normal Trade Relations status to estimate the importance of constituent interests and gauge how much information politicians had. Assuming perfect foresight of the shock would bias downward the role of constituent interests. We therefore adopt the moment inequality approach of Dickstein and Morales (2018), which delivers unbiased estimates under weaker informational assumptions and offers a way to test the information set of politicians. We find a moderate role of constituent interests, compared to ideology, and we cannot reject that politicians could easily predict the China Shock in the early 1990's, but not around 2000, when China started entering new sectors. In short, politicians had substantial information, but did not react strongly to it.

* Bombardini: UC Berkeley Haas School of Business, CEPR and NBER; Li: National University of Singapore; Trebbi: UC Berkeley Haas School of Business, CEPR and NBER. We would like to thank seminar participants at various institutions.

1 Introduction

The China Shock, the large surge in imports from China that started in the 1990's and that has made China one of the US main trading partners, has received broad attention in academia and policymaking. Although the early papers by Autor et al. (2016) and Pierce and Schott (2016) focused mainly of its employment effects, a large literature has expanded the analysis to social and health outcomes.[1] Although in hindsight China's entry in the US market seems like a preordained outcome, in a series of roll call votes during the 1990's US legislators were faced with the choice of allowing China to maintain its Normal Trade Relations (NTR) with the US and ultimately gain Permanent NTR, a prerequisite to obtain WTO membership.

In this paper we ask whether US politicians knew about the consequences of the China shock and how much the expected impact on their constituents affected their vote in favor or against China's NTR status. These two questions are intrinsically related. An obvious approach to estimating the importance of constituent interests is to replace the politician's expectation of the district-level China Shock with its realized value in a simple discrete choice voting model. However, this assumption that politicians are perfectly informed about the future shock implies necessarily a downward bias in the coefficient that measures the importance of constituent interests (unless politicians truly have perfect foresight). A small coefficient would be interpreted as low responsiveness of politicians to their constituents, indicating an accountability problem. Instead, a small coefficient could simply be the result of assuming that politicians are better informed than they truly are, pointing to limited expertise or insufficient funding for information acquisition.

We thus employ the moment inequality methodology of Dickstein and Morales (2018), developed in the context of firms' decision to export. This method allows us not only to estimate our voting model under much less stringent assumptions about the politicians information set, but it also offers tools to answer the question: how much did politicians know about the China Shock?

We employ the 11 yearly roll call votes between 1990 and 2001 to estimate a probabilistic voting model that depends on ideology and the China Shock. The first result is a moderate role of constituent interests. An interquartile difference in the value of the the shock increases the probability of voting in favor of NTR for China by roughly 4 percentage points, while a similar difference in ideology creates a 15-20 percentage point increase. In heterogeneity analysis, constituent interests are more important for Democrats than for Republicans and for politicians that were elected with small vote margins.

The second result is that politicians possess a significant amount of knowledge about of the future China Shock and in some years we cannot reject that they perfectly forecasted the shock that would hit their district in the next five years. Perhaps surprisingly, knowledge decreases

[1] Feler and Senses (2017),Dorn et al. (2019), Greenland and Lopresti (2016) and Greenland et al. (2019)

over the 1990's, a result that is easier to understand when we show that China's comparative advantage changed substantially during the late 1990's and early 2000's. Comparing legislators across parties, we find that Democrats are systematically more informed than Republicans, with the exception of the early 1990's.

Finally, we employ the estimated model to perform a counterfactual exercise in which we give politicians perfect information about the upcoming shocks and calculate the change in voting behavior that would have resulted from the additional information. Differently from Dickstein and Morales (2018), we find that overall support for China's NTR status would not have changed in the presence of perfect information.

One important premise of the question we are posing is the assumption that the electorate is generally attuned to trade policy positions of their representatives. While it is implausible to assume that voters have a command of specific trade policy measures, a number of papers have documented the impact of the China shock and the recent trade war on electoral outcomes. Autor et al. (2020) find that district more affected by the China shock saw an increase in Fox viewership and elected more conservative Republicans and, to a lesser extent, more liberal Democrats, thus inducing more polarization.[2] Another recent contribution by Che et al. (2016) finds that the 2010's reaction to the China shock is due to the anti-trade turn taken by the Republican party after the appearance of the Tea Party. The same paper finds that, during the 2000's, the areas affected by the China shock voted more in favor of Democrats. Blanchard et al. (2019) also document a significant electoral impact of the trade war on the vote share of Republicans in 2018. Interestingly, the negative effect on GOP vote share coming from retaliatory tariffs imposed by US trading partners is not mirrored by a positive effect due to the protection offered by import tariffs. In sum these recent papers offer a clear justification for making constituent interests a major component of the decision to vote on an important trade policy measure like maintaining and expanding China's Normal Trade Relations status.

We also contribute to a rich literature that links the structure of tariffs to political considerations, either through electoral support or to interest groups activity. A recent indication that constituent interests play an important role is given by Fajgelbaum et al. (2020)who show that US import tariffs in the recent trade war with China are such that pivotal counties (those with a Republican vote share of around 50%) receive the highest level of protection.

[TO BE COMPLETED]

[2]Colantone and Stanig (2018) document a similar result for Western European countries.

2 Empirical model

In this section we build a simple model of probabilistic voting for members of Congress. Indicate a congressional cycle with $t = 1, 2, ..., T$. Each period offers the possibility of introducing a single bill focused on a main policy issue - in this application the normalization of trade relations with China. As previously discussed, such bills were typically presented to the U.S. legislative branch and voted upon once per congressional cycle, so that t may equivalently indicate time and bill number. Let us indicate with $x_t \in \mathbb{R}$ a policy position favorable to trade normalization, so that and a Yes vote will indicate a vote for x_t. (Consequently, one interprets a No vote as a vote against normal trade relations, $q_t \in \mathbb{R}$.) This notation allows for the two positions to be affected by some nuance over time and neither position is assumed to be exactly constant in time. In the empirical analysis we will simply make sure that a Yes vote will be consistently labeled to indicate support for x_t.

Indicate by $i = 1, ..., N$ individual legislators, where N is large.[3] We will assume that individual i's preferences are represented by a random utility framework within a spatial voting environment. Spatial voting is a successful and informative modeling approach to the description of congressional behavior and it has found substantial support in the literature (Poole and Rosenthal 1997; Heckman and Snyder 1997; Clinton et al. 2004; McCarty et al. 2006; Bateman et al. 2017).

For simplicity of exposition (relaxed in the empirical application later), the deterministic component of the politician's utility is assumed to depend on: (i) the distance of the bill from his/her ideological position θ^i; (ii) an electoral motive, summarized by his/her expected future electoral support $V_{i,t+1}$ (for example, due to the employment consequences of i's voting behavior from the perspective of his/her constituents).

Concerning (i), political ideology $\theta_i \in \mathbb{R}$ is a unidimensional and fixed characteristic of i. The assumption of unidimensionality is appropriate in the time period under analysis (see McCarty et al. 2006). The assumption of constant policy preferences has been validated repeatedly in the literature on Congress (Moskowitz et al. 2017 for a discussion). We will use the ideological positions θ^i from DW-Nominate scores (Poole and Rosenthal, 1997).[4] This follows a common approach in modeling congressional voting when the explicit estimation of such preference parameters is peripheral to the main empirical investigation (Mian et al. 2010, 2014).[5]

Concerning (ii), let us indicate by $S_{i,t}$ a proxy for the degree of exposure of the local labor market to increasing imports from China (the "China shock", Autor et al. 2016) in the district

[3] What follows can be applied to each chamber independently at the cost of omitting interactions between the two chambers, such as resolutions and conferences. N =435 for the House and $N = 100$ for the Senate.

[4] For further reference, see www.voteview.com

[5] The reader interested in the estimation of θ^ican find a detailed analysis in Canen et al. (forth.) and references therein. Due to lack of sample overlap the Canen et al. (forth.) estimates cannot be used in our application.

represented by i at time t. Assume the potential electoral impact of the China shock in the district represented by politician i at time t is defined by:

$$V_{i,t+1} = h_t(d_t, S_{i,t+1}) + e_{i,t+1},$$

where d_t is the voting decision made by the politician, and

$$\begin{aligned} \mathbb{E}\left[e_{i,t+1}|d_t, S_{i,t+1}, \mathcal{I}_{i,t}\right] &= 0 \\ h_t(d_t, S_{i,t+1}) &= \gamma_t^0 + \gamma_t^1 S_{i,t+1} \times \mathbf{1}\left\{d_t = \text{vote for } x_t\right\} \\ &\quad + \gamma_t^2 S_{i,t+1} \times \mathbf{1}\left\{d_t = \text{vote for } q_t\right\} \\ &= \gamma_t^0 + \gamma_t^2 S_{i,t+1} + \left(\gamma_t^1 - \gamma_t^2\right) S_{i,t+1} \times \mathbf{1}\left\{d_t = \text{vote for } x_t\right\} \end{aligned} \tag{1}$$

where$\mathbf{1}\{.\}$ is an indicator function and $\mathcal{I}_{i,t}$ is the information set of politician i at t. Note that the function $h_t(.)$ introduces both a direct effect of the China shock on electoral support independently of i's vote and a component that depends on the interpretation by voters of their representative's decision. Such electoral effects are allowed to vary over time. Note that we introduce the China Shock directly in the electoral outcome for politician i, but such its effect can be interpreted as a composite two forces: the impact of the shock on employment outcomes and the impact of employment outcomes on electoral results.

The expected utility for a politician i of taking decision d_t given information set $\mathcal{I}_{it}$ is:

$$\begin{aligned} &U(\xi_t, d_t; \theta_i, \mathcal{I}_{i,t}) \\ &= u(d_t; \theta_i, \mathcal{I}_{i,t}) + \begin{cases} \xi_{i,t,x} & if\ d_t = \text{vote for } x_t \\ \xi_{i,t,q} & if\ d_t = \text{vote for } q_t \end{cases} \\ &= u\left(\parallel d_t - \theta^i \parallel\right) + \tilde{\delta}\mathbb{E}\left[\Sigma_{k=1}^K V_{i,t+k}(1+r)^{1-k}|d_t, \mathcal{I}_{i,t}\right] \\ &+ \begin{cases} \xi_{i,t,x} & if\ d_t = \text{vote for } x_t \\ \xi_{i,t,q} & if\ d_t = \text{vote for } q_t \end{cases}, \end{aligned}$$

where $u\left(\parallel . \parallel\right)$ indicates an ideological loss that is function of the distance of the policy from the ideal point of i. The term $\Sigma_{k=1}^K V_{i,t+k}(1+r)^{-k}$ indicates the $r-$discounted K future electoral outcomes for the district represented by i. For instance, $K = 1$ if politician i considers only the electoral impact of employment in the next congressional cycle election. For simplicity of exposition of the estimation approach, we assume $K = 1$ and expand the empirical analysis to $K > 1$. We assume a quadratic loss function $u(.)$ and i.i.d. Gaussian $\xi_{i,t,d} \sim N(0, \sigma_\xi^2)$. This implies the useful convolution $\xi_{it} = \xi_{i,t,q} - \xi_{i,t,x} \sim N(0, 2\sigma_\xi^2)$. A standard identification requirement in

discrete choice problems with Gaussian shocks (e.g. in Probit) further requires the normalization $2\sigma_{\xi}^2 = 1$, which we impose.

We define the variable Y_{it} as an indicator function that is equal to 1 when legislator i decides to vote Yes on x_t and zero when the legislator votes in favor of the status quo q_t :

$$\begin{aligned} Y_{it} &= \mathbf{1}\{U(\xi_t, x_t; \theta_i, \mathcal{I}_{i,t}) > U(\xi_t, q_t; \theta_i, \mathcal{I}_{i,t})\} = \\ &= \mathbf{1}\left\{u\left(\parallel x_t - \theta^i \parallel\right) - u\left(\parallel q_t - \theta^i \parallel\right) + \tilde{\delta}\left(\mathbb{E}\left[V_{i,t+1}|x_t, \mathcal{I}_{i,t}\right] - \mathbb{E}\left[V_{i,t+1}|q_t, \mathcal{I}_{i,t}\right]\right) > \xi_{it}\right\} = \\ &= \mathbf{1}\left\{-\frac{1}{2}\left(\left(x_t - \theta^i\right)^2 - \left(q_t - \theta^i\right)^2\right) + \tilde{\delta}\left(\mathbb{E}\left[V_{i,t+1}|x_t, \mathcal{I}_{i,t}\right] - \mathbb{E}\left[V_{i,t+1}|q_t, \mathcal{I}_{i,t}\right]\right) > \xi_{it}\right\}. \end{aligned} \tag{2}$$

We can write the probability of $Y_{i,t} = 1$ as:

$$\begin{aligned} &Pr(Y_{i,t} = 1|\mathcal{I}_{i,t}) \\ &= \Phi\left(\begin{array}{c} -\frac{1}{2}\left(\left(x_t - \theta^i\right)^2 - \left(q_t - \theta^i\right)^2\right) \\ +\tilde{\delta}\left(\mathbb{E}\left[V_{i,t+1}|x_t, \mathcal{I}_{i,t}\right] - \mathbb{E}\left[V_{i,t+1}|q_t, \mathcal{I}_{i,t}\right]\right) \end{array}\right), \end{aligned}$$

where Φ is the standard normal cumulative density function and $\mathbb{E}\left[V_{i,t+1}|x_t, \mathcal{I}_{i,t}\right] - \mathbb{E}\left[V_{i,t+1}|q_t, \mathcal{I}_{i,t}\right]$ is the expected net loss (or net gain) of electoral support due to a "China shock" vote in the constituency represented by politician i in the $K = 1$ future electoral cycles, given the information available to i at t. This implies that the probability of voting Yes depends on the expectations of the electoral consequences of voting Yes relative to voting No, which are unobserved by the econometrician. The higher is the expected electoral gain of voting Yes, the higher the likelihood of voting Yes.

It follows from (1) that:

$$\mathbb{E}\left[V_{i,t+1}|x_t, \mathcal{I}_{i,t}\right] - \mathbb{E}\left[V_{i,t+1}|q_t, \mathcal{I}_{i,t}\right] = \left(\gamma_t^1 - \gamma_t^2\right)\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right].$$

Setting $\delta_t = \tilde{\delta}\left(\gamma_t^1 - \gamma_t^2\right)$ and simplifying the relative loss function as $a_t\theta^i + b_t = -\frac{1}{2}\left(\left(x_t - \theta^i\right)^2 - \left(q_t - \theta^i\right)^2\right)$, where $a_t = x_t - q_t$ and $b_t = \frac{1}{2}(q_t^2 - x_t^2)$, we can further rewrite:

$$\begin{aligned} &Pr(Y_{i,t} = 1|\mathcal{I}_{i,t}) \\ &= \Phi\left(a_t\theta^i + b_t + \delta_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]\right). \end{aligned} \tag{3}$$

A key parameter is δ_t, which gauges the weight in the politician's vote choice of his/her expectations about the future China shock, $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]$. We discuss two distinct approaches to the estimation of this and the other parameters in the next section.

3 Expectations and information set of politicians

A key contribution of this paper is the analysis of the information set available to politicians to forecast the labor market effects of the China Shock at the time of a roll call vote. As emphasized by Dickstein and Morales (2018), there are two fundamentally different approaches, which in turn hinge on the answer to the following question: is the politician's information set $\mathcal{I}_{i,t}$ known to the econometrician? When the answer is in the affirmative, then estimation can be performed by maximum likelihood. When the econometrician knows only a subset of the information available to politicians, then we adopt the moment inequality estimator proposed by Dickstein and Morales (2018). We discuss these two approaches in turn, but we first start with describing the three possible information sets that we will consider throughout the paper:

(i) Minimal Information: the politician knows his own ideological position θ^i, but the only information a politician has about the economic impact of the China Shock is the current share of population employed in manufacturing

(ii) Baseline Information: the politician has access to the minimal information set and to the current period China shock S_{it}

(iii) Perfect Foresight: the politician has perfect foresight of the labor market consequences of the China Shock so that $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right] = S_{i,t+1}$.

3.1 Politician's information set fully known to the econometrician: MLE

When the econometrician knows the content of the politician's information set, then the parameter vector $\omega_t = \left\{x_t, q_t, \tilde{\delta}_t\right\}$ can be estimated by maximum likelihood for each cycle. Based on expression (3) the log-likelihood function takes the form:

$$\begin{aligned} & ln\mathcal{L}\left(\omega_t | \left\{Y_{i,t}, \theta^i, \mathcal{I}_{i,t}\right\}_{i=1}^N\right) \\ & = \sum_{i=1}^N Y_{i,t} ln\left[\Phi\left(a_t\theta^i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]\right)\right] \\ & + \left(1 - Y_{i,t}\right) ln\left[1 - \Phi\left(a_t\theta^i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]\right)\right]. \end{aligned} \tag{4}$$

Maximizing (4) requires specifying the information set $\mathcal{I}_{i,t}$. In the case of perfect foresight, (4) is maximized after replacing $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]$ with $S_{i,t+1}$. In the case of Minimal information set, the expectation of the China Shock is derived as the predicted value of the following OLS regression:

$S_{i,t+1} = \beta_0 + \beta_1 \theta^i + \beta_2 ShareMfg_{it} + \varepsilon_{i,t+1}$.[6] For the Baseline information set, we can perform a similar two-step procedure, albeit with an OLS regression that contains a larger set of regressors, reflecting a more precise knowledge by the politician.

The key assumption of the maximum likelihood approach is that we, as econometricians, are confident about what enters the politician's information set. As shown by Dickstein and Morales (2018), when we misspecify the politician's information set, our parameter estimates ω_t will be biased. The direction of the bias cannot be characterized in general, so we resort to Monte Carlo simulations to illustrate the problem in Appendix B, but one specific case lends itself to an intuitive explanation. When the econometrician incorrectly assumes that the politician has perfect foresight, the bias that arises is similar to the case of classical measurement error in a linear regression setting. The intuition is that $\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}]$ is measured with error and, similarly to a linear regression setting, this will lead to an attenuation bias in the estimated coefficient δ_t. Assume that the true δ is negative and consider two representatives in districts A and B, who form their expectations based only on a minimal information set, which includes the manufacturing share in the region. Assume that district A and B have similar manufacturing shares, but different industrial composition. Hence the two representatives predict a similar import shock, but in reality district A is much more severely affected than district B. We, the econometricians, assume that these representatives are instead very well informed about the imminent import increases. Because A and B expect a similar impact, they vote similarly on the bill. The econometrician, however, observing a similar voting behavior between politicians A and B, concludes that δ is smaller (in absolute value) and that the politicians place little weight on the import shock.

3.2 Politician's information set partially known to the econometrician: Moment inequality approach

In the previous section we have shown that the maximum likelihood approach relies on an accurate knowledge by the econometrician of the information set possessed by the politician. The alternative estimation method proposed by Dickstein and Morales (2018), based on moment inequalities, does not require full knowledge of $\mathcal{I}_{i,t}$, but rather of a subset of variables $Z_{i,t} \subseteq \mathcal{I}_{i,t}$. That is, the politician may know more than $Z_{i,t}$ in forming his/her forecast, but knows at least the covariates $Z_{i,t}$, which are then available instruments for the computation of i's expectations. Assuming only partial knowledge of $\mathcal{I}_{i,t}$ comes at the cost of identification. We will not be able to point identify the elements of the parameter vector ω_t, but only to set identify them. Whether the confidence sets we present are informative will be analyzed in detail.

The second important goal of the following analysis is to ascertain the extent of legislators'

[6] SeeManski (1991) and Ahn and Manski (1993)

information set, for which we will follow the identification and selection methodology proposed by Dickstein and Morales (2018). This involves a formal analysis of which subset of variables a politician considers at the time of his/her vote through an application of specification selection tests proposed by Bugni et al. (2015). Being able to reject that certain variables are used in the politician's forecast allows us to learn about the process of decision making of legislators in this particular setting: what they knew and considered relevant at the time of their forecasts. In this exercise the voting model and data are kept constant, but the subset of variables assumed part of the information set is varied.

We will allow politicians to have time varying information sets and we will formally test whether certain groups of legislators have identical information sets or not, by following Dickstein and Morales (2018). For instance, we assess whether members of higher levels of seniority have broader information sets than lower seniority members, or whether members of opposing parties share the same information set.

A final question that the approach will allow us to answer is whether, had politicians had a more complete information set, their votes for trade normalization with China would have been different. These counterfactuals are simulated within the same structure of expectations and information we just described.

3.2.1 Odds-based moment inequalities

Throughout, we maintain the assumption of rational expectations on the part of politicians, that is the expectational error $\varepsilon_{i,t+1} = S_{i,t+1} - \mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]$ has mean zero, $\mathbb{E}\left[\varepsilon_{i,t+1}|\mathcal{I}_{i,t}\right] = 0$. Intuitively, politicians do not systematically skew their prediction or ignore elements of their information set which would systematically help in forecasting $S_{i,t+1}$. We follow Dickstein and Morales (2018) in generating two sets of moment inequalities that identify the possible values that the parameters of interest can take. That is, we set identify ω_t applying the estimator developed in Pakes et al. (2015). The derivations follow closely Dickstein and Morales (2018), so we refer the reader to their Appendix C for details.

We use the definition in (2) to obtain:

$$\mathbf{1}\{a_t\theta^i + b_t + \delta_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right] - \xi_{it} \geq 0\} - Y_{it} = 0. \tag{5}$$

We cannot work directly with this expression, as it depends on the unobserved shock realization ξ_{it} and $\mathcal{I}_{i,t}$. Therefore, we take the expectation of (5) conditional on $\mathcal{I}_{i,t}$ and θ^i and manipulate the expression to obtain the following equality:

$$\mathbb{E}\left[(1 - Y_{it})\frac{\Phi\left(a_t\theta^i + b_t + \delta_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]\right)}{1 - \Phi\left(a_t\theta^i + b_t + \delta_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]\right)} - Y_{it}\middle|\mathcal{I}_{i,t}, \theta^i\right] = 0 \tag{6}$$

This equality still depends on the expectation $\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}]$, which in turn depends on the true information set $\mathcal{I}_{i,t}$, an object that we also dot not observe. However, under the assumption that the expectational error $\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}]$-$S_{i,t+1}$ has mean zero and from the property that $\frac{\Phi}{1-\Phi}$ is convex, one can replace $\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}]$ with $S_{i,t+1}$ and apply Jensen's inequality to derive the following inequality:

$$\mathbb{E}\left[(1-Y_{it})\frac{\Phi\left(a_t\theta^i+b_t+\delta_t S_{i,t+1}\right)}{1-\Phi\left(a_t\theta^i+b_t+\delta_t S_{i,t+1}\right)}-Y_{it}\middle|\mathcal{I}_{i,t},\theta^i\right]\geq 0 \tag{7}$$

Consider now a subset of the information set $Z_{it}\subseteq\mathcal{I}_{i,t}$. Because Z_{it} is a subset of $\mathcal{I}_{it}$, the distribution of Z_{it} conditional on $\mathcal{I}_{it}$ is degenerate. This is to say that once $\mathcal{I}_{it}$ is known, then Z_{it} conditional on $\mathcal{I}_{it}$ is deterministic.

Starting from 6 and by applying the Law of Iterated Expectations, we can derive the following equality:[7]

$$\mathbb{E}\left[(1-Y_{it})\frac{\Phi\left(a_t\theta^i+b_t+\delta_t\mathbb{E}[S_{i,t+1}|Z_{i,t}]\right)}{1-\Phi\left(a_t\theta^i+b_t+\delta_t\mathbb{E}[S_{i,t+1}|Z_{i,t}]\right)}-Y_{it}\middle|Z_{i,t},\theta^i\right]=0$$

which, by the same arguments employed to obtain (7), yields the inequality:

$$\mathbb{E}\left[m_l^{ob}\middle|Z_{i,t},\theta^i\right] \geq 0 \tag{8}$$

$$m_l^{ob} = (1-Y_{it})\frac{\Phi\left(a_t\theta^i+b_t+\delta_t S_{i,t+1}\right)}{1-\Phi\left(a_t\theta^i+b_t+\delta_t S_{i,t+1}\right)}-Y_{it} \tag{9}$$

Notice that (9) is increasing in $\tilde{\delta}_t$ so it identifies a lower bound for this parameter.

Following a similar logic, one can derive a moment condition that further bounds the parameters of interest $\tilde{\delta}_t$, x_t and q_t:

$$\mathbb{E}\left[m_u^{ob}\middle|Z_{i,t},\theta^i\right] \geq 0 \tag{10}$$

$$m_u^{ob} = Y_{it}\frac{1-\Phi\left(a_t\theta^i+b_t+\delta_t S_{i,t+1}\right)}{\Phi\left(a_t\theta^i+b_t+\delta_t S_{i,t+1}\right)}-1+Y_{it} \tag{11}$$

Notice that m_u^{ob} is decreasing in $\tilde{\delta}_t$ and therefore moment inequality (10) identifies an upper bound for this parameter.

[7]Equation (6), which is of the form $\mathbb{E}[X|\mathcal{I}]=0$, can be rewritten as $\mathbb{E}_z[\mathbb{E}_x[X|Z]|\mathcal{I}]=\int_z\mathbb{E}_x[X|z]f_z(z|\mathcal{I})dz=0$. The last equality follows from the fact that, conditional on $\mathcal{I}$, Z has a degenerate distribution, i.e. $f_z(z|\mathcal{I})=1$.

3.3 Revealed preference moment inequalities

The second set of moment inequalities derives from the revealed preference argument that a politician will vote Yes if and only if the benefit from doing so is positive, hence:

$$Y_{it}\left[a_t\theta^i + b_t + \tilde{\delta}_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right] - \xi_{it}\right] \geq 0 \tag{12}$$

Because, again, ξ_{it} is unobserved, we take the expectation of (12), conditional on θ^i and $\mathcal{I}_{i,t}$ and obtain the following inequality:

$$Y_{it}\left[a_t\theta^i + b_t + \tilde{\delta}_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]\right] + \Gamma_{it} \geq 0 \tag{13}$$

where $\Gamma_{it} = -\mathbb{E}\left[Y_{it}\xi_{it}|\,\theta^i, \mathcal{I}_{i,t}\right] = (1-Y_{it})\frac{\phi\left(a_t\theta^i+b_t+\tilde{\delta}_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}]\right)}{1-\Phi\left(a_t\theta^i+b_t+\tilde{\delta}_t\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}]\right)}$. Once again the expression in inequality (13) contains the unobserved expectation $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]$. Because $\frac{\phi}{1-\Phi}$ is convex, we can replace the expectation with $S_{i,t+1}$ and take the expectation conditional on Z_{it}, the subset of $\mathcal{I}_{i,t}$. The resulting inequality will be weaker than (13) and is given by:

$$\begin{aligned} \mathbb{E}\left[m_l^{rp}|\,Z_{it}\right] &\geq 0 \\ m_l^{rp} &= Y_{it}\left[a_t\theta^i + b_t + \delta_t S_{i,t+1}\right] + (1-Y_{it})\frac{\phi\left(l\left(x_t,q_t,\theta^i\right)+\delta_t S_{i,t+1}\right)}{1-\Phi\left(l\left(x_t,q_t,\theta^i\right)+\delta_t S_{i,t+1}\right)} \end{aligned} \tag{14}$$

Starting from another revealed preference inequality:

$$(1-Y_{it})\left[\xi_{it} - a_t\theta^i - b_t + \delta_t\mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}\right]\right] \geq 0,$$

we can obtain a second revealed-preference moment inequality in a similar manner:

$$\begin{aligned} \mathbb{E}\left[m_u^{rp}|\,Z_{it}\right] &\geq 0 \\ m_u^{rp} &= -(1-Y_{it})\left[a_t\theta^i + b_t + \delta_t S_{i,t+1}\right] + Y_{it}\frac{\phi\left(a_t\theta^i + b_t + \delta_t S_{i,t+1}\right)}{\Phi\left(a_t\theta^i + b_t + \delta_t S_{i,t+1}\right)}. \end{aligned} \tag{15}$$

The moment inequalities defined by (8), (10), (14) and (15) are conditional on values of the vectors Z_{it}, which we allow to contain different variables. In particular, we allow Z_{it} to coincide with the minimal information set, i.e. $Z_{it} = \{\theta^i, ShareMfg_{it}\}$ or to include the current China shock S_{it}.

3.4 Moments for estimation

As explained in Dickstein and Morales (2018), conditional moments (8), (10), (14) and (15) cannot be directly employed for empirical applications because conditioning on each possible value of Z_{it} is

unfeasible. Their solution, which we adopt, is to transform conditional into unconditional moment inequalities.[8] This is not innocuous in that information is lost in transitioning from conditional to a relatively smaller set of unconditional inequalities. We return to the issue of whether this induces confidence sets that are too large when we present our results.

We collect the four moment inequalities (8), (10), (14) and (15) and, as in Dickstein and Morales (2018), we adopt the unconditional moment inequalities:

$$\mathbb{E}\left[\left\{\begin{array}{c} m_l^{ob} \\ m_u^{ob} \\ m_l^{rp} \\ m_u^{rp} \end{array}\right\} \times g\left(Z_{it}\right)\right] \geq 0$$

where the instrument function $g\left(Z_{it}\right)$ is specified, for the case of $Z_{it} = \{\theta^i, ShareMfg_{it}, S_{it}\}$, as follows:

$$g_a\left(Z_{it}\right) = \left\{\begin{array}{c} \mathbf{1}\left\{\theta_i > med\left(\theta_i\right)\right\} \times \left(\left|\theta_i - med\left(\theta_i\right)\right|\right)^a \\ \mathbf{1}\left\{\theta_i \leq med\left(\theta_i\right)\right\} \times \left(\left|\theta_i - med\left(\theta_i\right)\right|\right)^a \\ \mathbf{1}\left\{ShareMfg_{it} > med\left(ShareMfg_{it}\right)\right\} \times \left(\left|ShareMfg_{it} - med\left(ShareMfg_{it}\right)\right|\right)^a \\ \mathbf{1}\left\{ShareMfg_{it} \leq med\left(ShareMfg_{it}\right)\right\} \times \left(\left|ShareMfg_{it} - med\left(ShareMfg_{it}\right)\right|\right)^a \\ \mathbf{1}\left\{S_{it} > med\left(S_{it}\right)\right\} \times \left(\left|S_{it} - med\left(S_{it}\right)\right|\right)^a \\ \mathbf{1}\left\{S_{it} \leq med\left(S_{it}\right)\right\} \times \left(\left|S_{it} - med\left(S_{it}\right)\right|\right)^a \end{array}\right.$$

for $a \in \{0, 1\}$.

[TO BE COMPLETED]

4 Data

4.1 Roll call votes

The data for roll call votes in the House of Representatives is collected from voteview.com. We employ 12 roll call votes from 1990 to 2001, as listed in table 1. These were House Joint Resolutions which sought to prevent the President from extending China's NTR status. These resolutions did pass in 9 out of the 12 votes and passed in 3 cases in 1990, 1991 and 1992, but no action was taken in the Senate.

[8] Starting from a moment inequality of the form $E[m|Z] \geq 0$, let us consider an instrument function $g\left(Z\right) > 0$. Multiplying the two yields $g\left(Z\right) E[m|Z] \geq 0$ which implies $E\left[g\left(Z\right) m\right] \geq 0$ whenever $g\left(Z\right)$ is Z-measurable.

4.2 The "China shock"

Our empirical analysis evaluates the the impacts of expected exposure to future "China shock" on the voting outcomes of the NTR bills at the congressional district level. The exposure faced by the congressional representative in a given district are generated from the import shocks in different local labor markets nested within the district. We start with constructing the import shocks at the level of commuting zones (CZs), which are clusters of adjoining counties characterized by strong commuting ties and have been conceptualized as local labor markets in the literature (Autor et al., 2013; Acemoglu et al., 2016). The shocks from different CZs are then aggregated to the congressional districts (CDs).

4.2.1 Exposure to import shock at the commuting zone level

Future supply shocks from China faced by the commuting zone j is constructed according to:

$$S_{jt+1} = \sum_k \frac{L_{jk,t}}{L_{j,t}} \frac{\Delta M^{oth}_{kt+1}}{Y_{k,t} + M_{k,t} - X_{k,t}}. \tag{16}$$

In this expression, ΔM^{oth}_{kt+1} is the change in import of good k from China by eight other (non-US) high-income countries over 5 years in the future.[9] It reflects the rising supply capacity of China due to its economic reforms, and is arguably exogeneous to the U.S. product-demand shocks from the perspectives of the US local economies (**????**). The future import growth is then normalized by the contemporaneous absorption (U.S. industry output plus net imports, $Y_{k,t} + M_{k,t} - X_{k,t}$), which captures the supply component of the growth of Chinese import penetration at the industry level. $L_{jk,t}/L_{j,t}$ denote the share of industry k in CZ j's total employment in the period t.

The Bartik-style measure (16) summarizes the exposure of CZ j to China's future supply shocks from the standpoint of t. Having a prefect foresight of (16) not only requires the information on contemporaneous employment composition of the local labor market and domestic absorptions of different industries, but also knowledge on supply shocks from China 5 years in the future. In our analysis, the future shocks correspond to the import supply growth over the period 1990-1995, 1991-1996, ..., 2001-2006, which covers China's post-WTO-accession period when the US witnessed the most intense increase in import competition from China.

While the congressional representatives may not have full knowledge of future import shock as in (16), they may use the information of the past shocks to form expectation. For years 1993-2001,

[9] The eight other high-income countries are Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland.

we construct import shock in the past 5 years analogously as follows:

$$S_{jt} = \sum_k \frac{L_{ik,t-5}}{L_{i,t-5}} \frac{\Delta M_{kt}^{oth}}{Y_{k,t-5} + M_{k,t-5} - X_{k,t-5}}, \tag{17}$$

where ΔM_{kt}^{oth} denotes the change in import of good k from China by eight other (non-US) high-income countries over the past 5 years. Due to data constraints, for years 1990-1992, we use the 2-year-lagged variables to construct the past shocks.[10]

The baseline measures (16) and (17) follow the specification in Acemoglu et al. (2016) and Autor et al. (2020), and can be derived from the workhorse trade models with a gravity structure. However, different from the literature that focuses the the impacts of contemporaneous trade shocks on local economies, our study aims at evaluating the extent to which politicians foresaw the future import shock from China as in (16), and whether they acted on the relevant information when setting China-specific trade policies.

4.2.2 Trade, employment and output Data

This subsection describes the data sources that we employ to construct the commuting-zone-level measures (16) and (17).

Data on bilateral trade flows over the period 1988-2006 for the 4-digit Standard International Trade Classification (SITC) industries are obtained from UN Comtrade Database. We concord these data to four-digit Standard Industry Classification (SIC) industries.[11] Following Autor et al. (2013) and Acemoglu et al. (2016), we aggregate together a few four-digit industries to

[10] To be specific, for $t = 1990, 1991, 1992$

$$S_{it} = \sum_k \frac{L_{ik,t-2}}{L_{i,t-2}} \frac{\Delta M_{kt}^{oth}}{Y_{k,t-2} + M_{k,t-2} - X_{k,t-2}},$$

where ΔM_{kt}^{oth} denotes the change in import of good k from China by eight other (non-US) high-income countries over the past 2 years. As is discussed below, we use the data from County Business Patterns (CBP) to construct employment shares. For 1990-1992, calculating L_{ikt-5}/L_{it-5} requires the 1985-1987 CBP data with industries classified based on 1977 SIC codes. For the purpose of analysis, the data needs to be mapped to the 1987 SIC codes. However, the crosswalk from 1977 SIC to 1987 SIC involves many splits of industries. As a result, the concordance leads to a structural break in the employment measures for some localities over 1987-1998. For the concern of systematic measurement errors, we don't use the CBP data prior to 1988 for the main analysis.

[11] The crosswalk that cross-matches the four-digit SITC (Rev.2) industries and the four-digit SIC (1987 version) is constructed as follows. (1) We first map the four-digit SITC industries to the corresponding six-digit Harmonized Commodity Description and Coding System (HS) products based on the concordance provided by UN WITS (https://wits.worldbank.org/product_concordance.html). (2) We then apply the crosswalk from **?**, which assigns 6-digit HS products to 4-digit SIC industries. (3) Lastly, the four-digit SITC codes are then cross-matched with the four-digit SIC codes based on their relations with the six-digit HS codes.

ensure compatibility with the additional data on employment discussed below. Our final data set contains 397 manufacturing industries. We complement the trade data by the output data obtained from the NBER-CES data. All import and output amounts are inflated to 2007 US dollars using the Personal Consumption Expenditure (PCE) deflator.

Information on industry employment structure by CZs over 1988-2001 is derived from the County Business Patterns (CBP) data published by the US Census Bureau. The CBP tracks employment, firm size distribution, and payroll by county and industry annually. To protect confidentiality, employment for county-industry cells is sometimes reported as an interval instead of exact count. We use the fixed-point imputation algorithm developed by Autor et al. (2013) to derive employment for each county-industry cell. The county-level data is then aggregated to commuting zones using the concordances provided by Autor et al. (2013).[12]

4.2.3 Exposure to Import Shock at the Congressional District Level

Following Autor et al. (2020), we map economic outcomes in CZs to CDs as follows. We start with the geographic relationships between counties and congressional districts provided by the Missouri Census Data Center (MCDC).[13] Counties are sometimes split across different CDs, and the MCDC concordance provides information on the distribution of the county population in each CD. We then ascribe to each county-by-congressional district cell the CZ-level import shock that corresponds to the county, and weight each cell by its share of population in the district. Lastly, we aggregate the weighted shocks across cells to the CD level. By construction, if a district span multiple CZs, its exposure to China's rising import competition is the population-share-weighted average of the import shocks in these CZs. Since congressional districts have more or less equal population size, their have roughly the same weight in our analysis. In the empirical analysis, we denote the future and past import shocks at the CD level by S_{it+1} and S_{it}, respectively.

During the sample period 1990-2001, the boundaries of county-by-congressional district cells experienced a major change in 1993 but remained stable afterwards. Therefore, for 1990-1992, we map the CZ-level import shocks to congressional districts as defined for the 102^{nd} Congress. For 1993-2001, the mapping is based on the configuration of the 103^{rd} Congress. This treatment won't affect the consistency of our baseline analysis, because as is discussed below, we conduct the

[12]Industry classifications in CBP changed periodically — over 1988-1997, employment is classified using the SIC (1987 version) codes, while employment thereafter is expressed according to the North American Industry Classification System (NAICS). Using the crosswalk in **?**, we correspond the post-1997 data to the four-digit SIC industries.

[13]http://mcdc.missouri.edu/applications/geocorr1990.html

estimation by periods based on presidential administrations, and none of the subsamples spans over 1992-1993.

Figure 1 shows the cross-district averages of past and future import shocks. The import shocks are always positive throughout the sample period, but the future shocks move less in tandem with the past shocks in the later years. For the moment inequality estimation discussed below, we detrend the import shocks. The corresponding distributions reported in Table 2 reveal a substantial heterogeneity in exposures across districts. For the periods 1997-2001, 1993-1996, and 1990-1992, the inter-quartile ranges of future shock are 0.154, 0.078 and 0.143, respectively.

Table 4 validates relevance of the instruments. Specifically, we find that both manufacturing employment share and past import shock have statistically significant effects on future import shock. On a related note, the R-squared is lower in the later period, suggesting that the variance of the expectational errors gets larger over time.

5 Results from congressional voting on trade relations with China

In this section we discuss the results of estimating of equation (3) using the three information sets: Minimal, Baseline and Perfect Foresight. Table 5 reports estimation results when splitting the votes by period: the first period 1990-1992 covers the George H.W. Bush presidency, the second period 1993-1996 covers the first Bill Clinton presidency and third period 1997-2001 (in the robustness section we drop the year 2001) covers the second Bill Clinton presidency. The table presents estimates for the parameters a,b and δ. We will mostly discuss results for the importance of constituency interests δ, but will gauge its magnitude in relation to the role of ideology a.

Consider the 1997-2001 results. Panel A reports estimation results under the assumption that the politicians' information set includes at least the manufacturing share in their district and their own ideology. Under this assumption we report a confidence set $[-1.975, -0.137]$, so that all values in the CS are negative, as we would expect if politicians are more likely to vote against China's NTR status if they expect their constituency to be exposed to a larger shock. The confidence set is quite similar ($[-1.505, -0.020]$) under the assumption of Baseline information, which includes also the current shock. We return in the sext setion to the fact that under perfect foresight we find an empty confidence set. In general, when non-empty, confidence sets include negative values for δ and are tight enough to be informative.

The magnitude of the parameter δ can be illustrated as follows. Considering two districts in 1997-2001, the probability of voting in favor of China's NTR status decreases by between 0.077 and

0.004 when the value of the China shock goes from its 25th percentile to its 75th percentile.[14]These numbers are 0.08 and 0.009 for 1993-1996 and 0.087 and 0.004 for 1990-1992. This is a moderate effect, especially considering the effect of ideology. If we perform the analogous exercise, we find that an interquartile range shift in ideology calculate at the mean China shock value produces an increase in the probability of voting in favor of China of between 0.132 and 0.215 in 1997-2001 and between 0.186 and 0.344 in 1990-1992.[15] In sum, these comparisons allows us to conclude that the effect of ideology is much larger than the effect of constituent interests. We will discuss differences in the parameters estimates between subsets of legislators after introducing specification tests that allow us to gauge the information possessed by politicians.

5.1 Information Set

A fundamental first step in the analysis of politicians' decisions is to assess the exact extent of their information sets at the moment of a vote. As discussed in Dickstein and Morales (2018), a valid statistical approach is presented in Bugni et al. (2015). The test proposed by Bugni et al. (2015) allows to reject the null hypothesis that there exists a value of ω_t within the identified set that can rationalizes the set of moment inequalities presented in Section3.4. Specifically, consider two alternative information sets $Z^1_{i,t}$ and $Z^2_{i,t}$. Then in the case there is no value of the parameter vector for which the set of moment inequalities hold given $Z^1_{i,t}$, but there are values of ω_t within the identified set that can rationalizes the set of moment inequalities given $Z^2_{i,t}$, then one can infer that $Z^1_{i,t} \nsubseteq \mathcal{I}_{i,t}$ and cannot reject $Z^2_{i,t} \subseteq \mathcal{I}_{i,t}$. As discussed in Dickstein and Morales (2018), the rejection of the null in this framework may also indicate misspecification of the original model, so simple rejection of $Z^1_{i,t}$ cannot exclude misspecification. However, failure to reject $Z^2_{i,t}$ eliminates this second interpretation of the test.

Following the appendix of Dickstein and Morales (2018), we first introduce some notations. The test statistics for moment inequalities is

$$Q(\delta_p) = \sum_{k=1}^{K} \left(\min\{ \frac{\bar{m}_k(\delta_p)}{\hat{\sigma}_k(\delta_p)}, 0 \} \right)^2 \tag{18}$$

where δ_p is a point in the space Δ_g. Denote δ^* as the true parameter. The null hypothesis to test is $H_0 : \delta^* = \delta_p$. Following the Generalized Moment Selection (GMS) test in Andrews and Soares (2010), we simulate the asymptotic distribution of $Q(\delta_p)$. Specifically, the simulation is based on R draws from the multivariate normal distribution $N(0_K, I_K)$, and each draw ζ_r yields

[14]In particular, we calculate these percentage points as $\Phi\left(b + \delta S^{75th}_{t+1}\right) - \Phi\left(b + \delta S^{25th}_{t+1}\right)$ at the mean $\theta^i = 0$ assuming the two values of δ at the boundaries of the confidence set.

[15]For the period 1993-1996 we obtain a confidence set that contains zero.

the following statistics:

$$Q_r^{AA}(\delta_p) = \sum_{k=1}^{K} \left(\left(\min\{[\hat{\Omega}^{\frac{1}{2}}(\delta_p)\zeta_r]_k, 0\} \right)^2 \times \mathbf{1}\{\sqrt{n}\frac{\bar{m}_k(\delta_p)}{\hat{\sigma}_k(\delta_p)} \leq \sqrt{\ln n}\} \right) \tag{19}$$

where $\hat{\Omega}$ is the estimate of the variance-covariance matrix of $\frac{\bar{m}_k(\delta_p)}{\hat{\sigma}_k(\delta_p)}$. $[\hat{\Omega}^{\frac{1}{2}}(\delta_p)\zeta_r]_k$ is the kth element of the vector $\hat{\Omega}^{\frac{1}{2}}(\delta_p)\zeta_r$, and it is the simulated counterpart of $\frac{\bar{m}_k(\delta_p)}{\hat{\sigma}_k(\delta_p)}$ in equation (18). Pooling all Q_r^{AA} together renders the simulated distribution denoted by $Q^{AA}(\delta_p)$.

Let $\hat{c}^{AA}(\delta_p, 1-\alpha)$ denote the $(1-\alpha)$-quantile of $Q^{AA}(\delta_p)$. δ_p is included in the $(1-\alpha)\%$ confidence set if $Q(\delta_p) \leq \hat{c}^{AA}(\delta_p, 1-\alpha)$. Details are provided in appendix A.7.1 of Dickstein and Morales (2018).

P-value of the Test BP. For a given value of λ, we infer whether $\forall\ \delta_p \in \Delta_g$, $Q(\delta_p) > \hat{c}^{AA}(\delta_p, 1-\lambda)$. If so, we lower λ by a small amount, and repeat the procedure until reaching the value λ^{BP} such that $\exists\ \delta_p \in \Delta_g$ such that $Q(\delta_p) \leq \hat{c}^{AA}(\delta_p, 1-\lambda^{BP})$. λ^{BP} is then the p-value for the test BP.[16] As pointed out by Bugni et al. (2015), this test is too conservative. Therefore, we turn to the test RC and the test RS proposed by the authors as follows.

P-value of the Test RC. To conduct the test RC, we first calculate the minimum of test statistics (18) across all $\delta_p \in \Delta_g$. Denote the minimum by $T = \min_{\delta \in \Delta_g} Q(\delta_p)$, and $\Delta_g^* = \arg\min_{\delta_p}(Q(\delta_p))$.[17] The p-value of the test RC is constructed as follows. For a given value of λ, we compute $\hat{c}^{RC}(1-\lambda) = \min_{\delta_p \in \Delta_g^*} \hat{c}^{AA}(\delta_p, 1-\lambda)$. If $T > \hat{c}^{RC}(1-\lambda)$, we lower λ by a small amount, and repeat the procedure until reaching the value λ^{RC} such that $T \leq \hat{c}^{RC}(1-\lambda^{RC})$. Then, λ^{RC} is the corresponding p-value.

P-value of the Test RS. Similar to the test RC, to conduct the test RS, we first compute T and derive Δ_g^* as defined above. Then, we use the R draws from the multivariate normal distribution $N(0_K, I_K)$. For each draw ζ_r, we compute $Q_r^{AA}(\theta_p)$ for each $\theta_p \in \Theta_g^*$ according to equation (19), and derive the corresponding minimum $T_r = \min_{\theta_p \in \Theta_g^*} Q_r^{AA}(\theta_p)$. The p-value of the test RS is constructed as follows. For a given value of λ, we find the $(1-\lambda)$-quantile of T_r, denoted by $\hat{c}^{RS}(1-\lambda)$. If $T > \hat{c}^{RS}(1-\lambda)$, we lower λ by a small amount, and repeat the procedure until reaching the value δ^{RS} such that $T \leq \hat{c}^{RS}(1-\delta^{RS})$. Then, δ^{RS} is the corresponding p-value.

Note that the test RC and the test RS are different in the following way. For the test RC, to

[16] In practice, we start the algorithm with $\delta = 0.99$ and reduce it by 0.005 at a time. We stop when δ reaches 0.01. Hence, p-value=0.01 (respectively, 0.99) in our tables indicates that the p-value is less than or equal to 0.01 (respectively, greater than or equal to 0.99).

[17] As discussed in Ho and Pakes (2014) (pg. 3868), Δ_g^* could be a set of values all of which make all the moments zero, or it could be a point, indicating that no value of δ satisfies all the moment conditions. The latter case could be a result of sampling error, which is accounted for by the GMS approach proposed by Andrews and Soares (2010). In our case, Θ_g^* contains only one point as the case in Ho and Pakes (2014).

derive the critical value $\hat{c}^{RC}(1-\lambda)$, we first compute the $(1-\lambda)$-quantile for each asymptotic distributions Q^{AA} evaluated at each $\delta_p \in \Delta_g^*$, and then take the minimum across these quantiles. For the RS test, to derive the critical value $\hat{c}^{RS}(1-\lambda)$, we first compute the minimum of $Q_r^{AA}(\delta_p)$ across $\delta_p \in \Delta_g^*$, and then derive the $(1-\lambda)$-quantile of these minimums. When Δ_g^* contains only one point, the $\hat{c}^{RC}(1-\lambda) = \hat{c}^{RS}(1-\lambda)$, and the p-values from the two tests are the same. When Δ_g^* contains multiple points, $\hat{c}^{RC}(1-\lambda) \geq \hat{c}^{RS}(1-\lambda)$, and the p-value of the test RS will be no larger than that of the test RC.

We now return to Table 5 to discuss columns 4-6, which report p-values for the three tests just described. For the full sample and for two of the three periods we reject that politicians have perfect foresight, with p-values of 0.01-0.015 depending on the test. For all periods we cannot reject that the politicians had at least the Baseline information set. Why would politicians be better informed in the earlier part of the period? We hypothesize that this is due to the more predictable nature of the shocks in the early 90's, when China was specialized in relatively less sophisticated products. In figure 2 we report the autocorrelation of the China shock and it is clear that in the earlier years the shock was more predictable from year to year as the autocorrelation was above 0.3. In the second half od the 90's the autocorrelation progressively dropped to zero, a fact that could justify why politicians were less than perfectly informed as the 2000's approached.

5.2 Heterogeneity Across Groups of Politicians

The estimates and tests presented so far were performed on the universe of the Members of the House during the 1990-2001 period. It is plausible to hypothesize that politicians from different parties, and with different electoral prospects, would have varying degrees of knowledge. In this section we analyze three dimensions of heterogeneity: political party, tenure and closeness of the most recent election. For all tests we report the Baseline information results. In table 6 we find that Democrats are more informed than Republicans as we cannot reject that Democrats had at least the Baseline information set in all subperiods, but we can reject the Baseline information set for Republicans in all periods but 1990-1992. When we obtain non-empty parameter estimates, we find that Democrats display higher (in absolute value) sensitivity to the China shock. Republicans' confidence set for δ straddles zero, while the entire confidence set for the Democratic members of the House is comprised of negative values. This is in line with Democratic legislators historical alignment with workers' interests.

In Table 7 we divide the sample in two according to whether House members tenure is above or below the median. One may anticipate that politicians that are very experienced have better access to various sources of information. We do not find support for this conjecture as confidence sets for δ of junior legislators are not different from those of senior ones.

In Table 8 we explore whether politicians with above-median victory margins in the previous election display differences in the sensitivity to constituent interests. The literature has found higher sensitivity for legislators in tighter races. We indeed find that, across different periods, legislators in tigher races display confidence sets for δ that are entirely composed of negative values, whereas confidence sets for politicians in safe races often include both positive and negative values.

Overall, the picture that emerges from the heterogeneity analysis is tht Democrats are both more informed and more sensitive to constituent interests, and legislators in close races respond more strongly to the China shock, but are not differentially informed relative to politicians elected with large margins.

5.3 Counterfactual: The Case of Legislators Informed About the China Shock

After establishing that, for most of the time period in our sample, politicians had less than perfect knowledge of the labor market consequences of China's export expansion, a natural question to ask is whether providing more accurate information to politicians would have changed their vote, and the overall passage of certain bills. The answer, in our case, is not much. Before we delve into the specific results, it is worth pointing out that when the expectational error is symmetrically distributed, there is no reason to anticipate an overall change in the share of Yes/No votes. This is because our assumption of rational expectations dictate that politicians are, on average, correct and so, while providing information can sway an individual politician to change her/his position, in the aggregate we will not see a substantial change in either direction. Figure **??** confirms that the expectational error is symmetric in this sample. It is worth pointing out that Dickstein and Morales (2018) find an aggregate change in the number of exporters because their data features highly asymmetric expectational errors (a few ferms make very large expectational errors). When this is the case, providing information can sway a large number of agents that were making a small expectational error (balanced by one large expectational error in the opposite direction).

Conditional on a particular value of the parameter vector $\omega_t = \{a_t, b_t, \delta_t\}$, and information set $\mathcal{I}_{it}$, the decision of voting in favor of China is given by:

$$y_{it}(\omega_t, \mathcal{I}_{it}, \xi_{it}) = \{a_t\theta_i + b_t + \delta_t \mathbb{E}[\Delta S_{i,t+1}|\mathcal{I}_{it}] - \xi_{it} \geq 0\}.$$

Integrating $y_{it}(\omega_t, \mathcal{I}_{it}, \xi_{it})$ over ξ_{it} generates the probability that politician i casts a pro-China vote in period t. In particular,

$$\int_{\xi_{it}} y_{it}(\omega_t, \mathcal{I}_{it}, \xi_{it})\phi(\xi_{it})d\xi_{it} = \int_{\xi_{it}} \mathbb{1}\{a_t\theta_i + b_t + \delta_t \mathbb{E}[\Delta S_{i,t+1}|\mathcal{I}_{it}] - \xi_{it} \geq 0\}\phi(\xi_{it})d\xi_{it}.$$

Denote $\mathcal{N}$ be the set of politicians, the total number of votes in favor of China is then given by:

$$N^{+}(\omega_t, \mathcal{I}_{it}, \mathcal{N}) = \sum_{i \in \mathcal{N}} \int_{\xi_{it}} y_{it}(\omega_t, \mathcal{I}_{it}, \xi_{it}) \phi(\xi_{it}) d\xi_{it}.$$

where $\Omega_i^{95\%}$ is the 95% confidence set for the underlying parameters.

The corresponding 95% confidence set of the number of politicians vote in favor of China is:

$$\left[\min_{\omega_t \in \Omega_t^{95\%}} \left\{ N^{+}(\omega_t, \mathcal{I}_{it}, \mathcal{N}) \right\}, \max_{\omega_t \in \Omega_t^{95\%}} \left\{ N^{+}(\omega_t, \mathcal{I}_{it}, \mathcal{N}) \right\} \right]$$

We simulate $N^{+}(\omega_t, \mathcal{I}_{it}, \mathcal{N})$ under the case of baseline information $\mathcal{I}_{it}^b$ and the case of perfect foresight $\mathcal{I}_{it}^p$. The percentage change in number of votes in favor of China when we provide politicians with full information is:

$$\left[\min_{\omega_t \in \Omega_t^{95\%}} \left\{ \frac{N^{+}(\omega_t, \mathcal{I}_{it}^p, \mathcal{N})}{N^{+}(\omega_t, \mathcal{I}_{it}^b, \mathcal{N})} - 1 \right\} \times 100, \max_{\omega_t \in \Omega_t^{95\%}} \left\{ \frac{N^{+}(\omega_t, \mathcal{I}_{it}^p, \mathcal{N})}{N^{+}(\omega_t, \mathcal{I}_{it}^b, \mathcal{N})} - 1 \right\} \times 100 \right]$$

The number of politicians who vote for China in the baseline but switch vote in the counterfactual is

$$\begin{aligned}
& N^{+-}(\omega_t, \mathcal{I}_{it}^b \to \mathcal{I}_{it}^p, \mathcal{N}) \\
&= \int_{\xi_{it}} y_{it}(\omega_t, \mathcal{I}_{it}^b, \xi_{it})(1 - y_{it}(\omega_t, \mathcal{I}_{it}^p, \xi_{it})) \phi(\xi_{it}) d\xi_{it} \\
&= \int_{\xi_{it}} 1\left\{ a_t \theta_i + b_t + \delta_t \mathbb{E}[\Delta M_{i,t+1} | \mathcal{I}_{it}^b] - \xi_{it} \geq 0 \right\} 1\left\{ a_t \theta_i + b_t + \delta_t \mathbb{E}[\Delta M_{i,t+1} | \mathcal{I}_{it}^p] - \xi_{it} < 0 \right\} \phi(\xi_{it}) d\xi_{it}.
\end{aligned}$$

Analogously, we can derive number of politicians in the other categories. Table 9 shows the results of this counterfactual exercise. There a few small changes in the voting patterns across politicians who now possess full information, a result of the fact that they already possess a substantial amount information. Moreove, the vote share in favor of China's NTR status is not different, because as discussed above, politicians are on average correct due to the assumption of rational expectations.

6 Concluding remarks

To be completed.

References

Acemoglu, D., Autor, D., Dorn, D., Hanson, G. H., Price, B., 2016. Import competition and the great us employment sag of the 2000s. Journal of Labor Economics 34 (S1), S141–S198.

Ahn, H., Manski, C. F., 1993. Distribution theory for the analysis of binary choice under uncertainty with nonparametric estimation of expectations. Journal of Econometrics 56 (3), 291–321.

Andrews, D. W., Soares, G., 2010. Inference for parameters defined by moment inequalities using generalized moment selection. Econometrica 78 (1), 119–157.

Autor, D., Dorn, D., Hanson, G., Majlesi, K., 2020. Importing political polarization? the electoral consequences of rising trade exposure. American Economic Review 110 (10), 3139–83.

Autor, D. H., Dorn, D., Hanson, G. H., 2013. The china syndrome: Local labor market effects of import competition in the united states. American Economic Review 103 (6), 2121–68.

Autor, D. H., Dorn, D., Hanson, G. H., 2016. The china shock: Learning from labor-market adjustment to large changes in trade. Annual Review of Economics 8, 205–240.

Bateman, D. A., Clinton, J., Lapinski, J. S., 2017. A house divided? political conflict and polarization in the u.s. congress, 1877- 2011. American Journal of Political Science 61 (3), 698–714.

Blanchard, E. J., Bown, C. P., Chor, D., 2019. Did trump's trade war impact the 2018 election? Tech. rep., National Bureau of Economic Research.

Bugni, F. A., Canay, I. A., Shi, X., 2015. Specification tests for partially identified models defined by moment inequalities. Journal of Econometrics 185 (1), 259–282.

Canen, N., Kendall, C., Trebbi, F., forth. Unbundling polarization. Econometrica.

Che, Y., Lu, Y., Pierce, J. R., Schott, P. K., Tao, Z., 2016. Does trade liberalization with china influence us elections? Tech. rep., National Bureau of Economic Research.

Clinton, J., Jackman, S., Rivers, D., 2004. The statistical analysis of roll call data. American Political Science Review 98 (2), 355–370.

Colantone, I., Stanig, P., 2018. The trade origins of economic nationalism: Import competition and voting behavior in western europe. American Journal of Political Science 62 (4), 936–953.

Dickstein, M. J., Morales, E., 2018. What do exporters know? The Quarterly Journal of Economics 133 (4), 1753–1801.

Dorn, D., Hanson, G., et al., 2019. When work disappears: Manufacturing decline and the falling marriage market value of young men. American Economic Review: Insights 1 (2), 161–78.

Fajgelbaum, P. D., Goldberg, P. K., Kennedy, P. J., Khandelwal, A. K., 2020. The return to protectionism. The Quarterly Journal of Economics 135 (1), 1–55.

Feler, L., Senses, M. Z., 2017. Trade shocks and the provision of local public goods. American Economic Journal: Economic Policy 9 (4), 101–43.

Greenland, A., Lopresti, J., 2016. Import exposure and human capital adjustment: Evidence from the us. Journal of International economics 100, 50–60.

Greenland, A., Lopresti, J., McHenry, P., 2019. Import competition and internal migration. Review of Economics and Statistics 101 (1), 44–59.

Heckman, J. J., Snyder, J. M., 1997. Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. The RAND Journal of Economics 28.

Ho, K., Pakes, A., 2014. Hospital choices, hospital prices, and financial incentives to physicians. American Economic Review 104 (12), 3841–84.

Manski, C., 1991. Nonparametric estimation of expectations in the analysis of discrete choice under uncertainty. In: nonparametric and Semiparametric methods in Econometrics and Statistics: proceedings of the fifth International Symposium in Economic Theory and Econometrics. Vol. 5. Cambridge University Press, p. 259.

McCarty, N., Poole, K. T., Rosenthal, H., 2006. Polarized America: The Dance of Ideology and Unequal Riches. Cambridge: MIT Press.

Mian, A., Sufi, A., Trebbi, F., December 2010. The political economy of the us mortgage default crisis. American Economic Review 100 (5).

Mian, A., Sufi, A., Trebbi, F., 04 2014. Resolving debt overhang: Political constraints in the aftermath of financial crises. American Economic Journal: Macroeconomics 6 (2), 1–28.

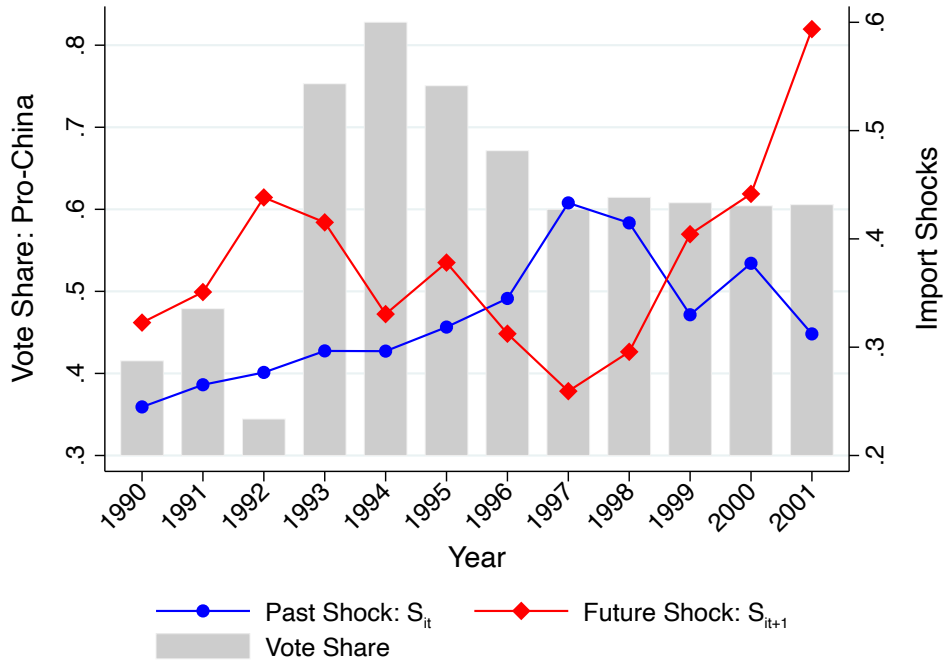
Moskowitz, D. J., Rogowski, J., Jr., J. M. S., 2017. Parsing party polarization, mimeo.

Pakes, A., Porter, J., Ho, K., Ishii, J., 2015. Moment inequalities and their application. Econometrica 83 (1), 315–334.

Pierce, J. R., Schott, P. K., 2016. The surprisingly swift decline of us manufacturing employment. American Economic Review 106 (7), 1632–62.

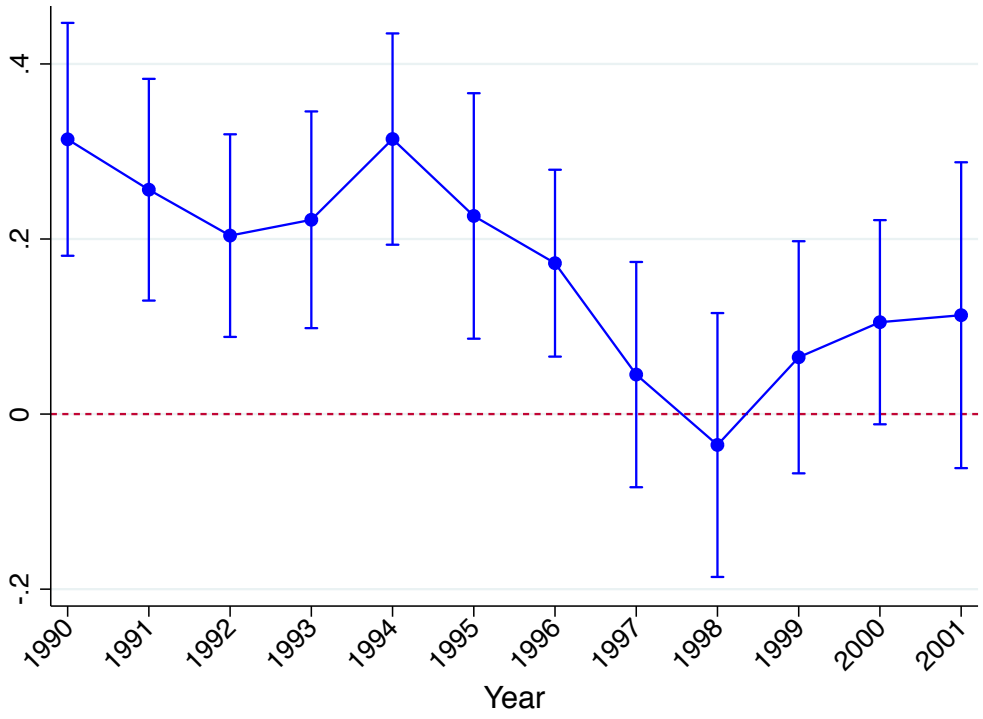
Poole, K. T., Rosenthal, H., 1997. Congress: A Political-Economic History of Roll Call Voting. New York: Oxford University Press.

Figure 1: Import Shocks from China and Pro-China NTR Vote Share



Notes: The bar chart shows the share of votes in favor of renewing Chinaâs NTR status for the bills introduced in House over 1990 to 2001. The data in 2000 also includes the bill HR 4444 which granted China permanent normal trade relations. The dot blue line represents the average of past import shock across congressional districts, and the diamond red line shows the average of the future import shock. In order to put the data on a comparable five-year scale, past import shocks over 1990-1992 are multiplied with the factor 5/2.

Figure 2: The China Shock over Time



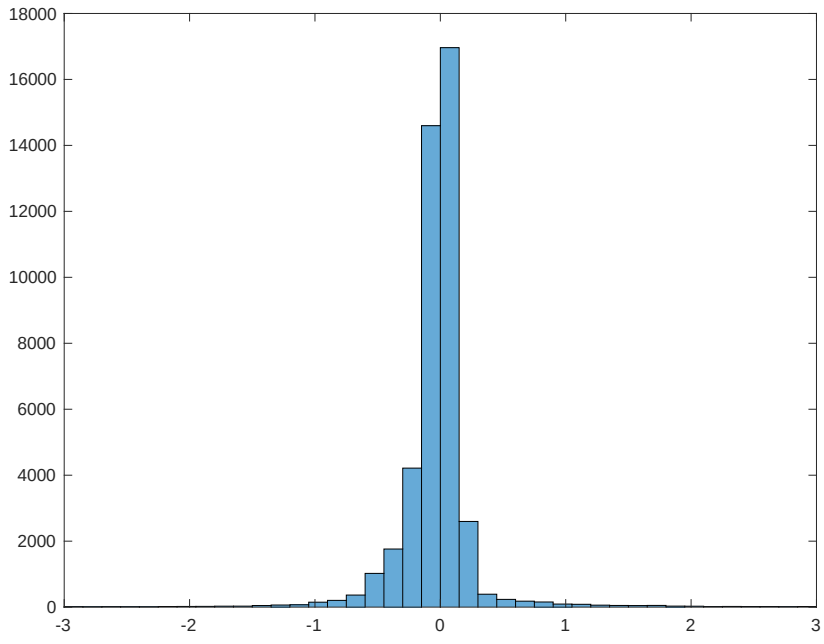


Figure 3: Distribution of Expecational Errors (DM) (Winsorized)

Table 1: Roll Call Votes

Year	Congress	President	House	Bill number	NTR approved in House	Additional action
1990	101	G.H.W. Bush	D	HJRES647	No	No action in Senate
1991	102	Bush Sr	D	HJRES263	No	No action in Senate
1992	102	Bush Sr	D	HJRES502	No	Did not pass in Senate
1993	103	Clinton	D	HJRES208	Yes	
1994	103	Clinton	D	HJRES373	Yes	
1995	104	Clinton	R	HJRES96	Yes	
1996	104	Clinton	R	HJRES182	Yes	
1997	105	Clinton	R	HJRES79	Yes	
1998	105	Clinton	R	HJRES121	Yes	
1999	106	Clinton	R	HJRES57	Yes	
2000	106	Clinton	R	HJRES103	Yes	
2001	107	G.W. Bush	R	HJRES50	Yes	

Table 2: Summary Statistics of Detrended Import Shocks at the Congressional District Level

	1990-2001		1997-2001		1993-1996		1990-1992	
	S_{it+1}	S_{it}	S_{it+1}	S_{it}	S_{it+1}	S_{it}	S_{it+1}	S_{it}
std	0.128	0.128	0.122	0.075	0.070	0.153	0.153	0.141
p5	-0.172	-0.175	-0.171	-0.106	-0.095	-0.186	-0.189	-0.187
p25	-0.077	-0.075	-0.084	-0.048	-0.043	-0.098	-0.093	-0.093
p50	-0.017	-0.016	-0.010	-0.009	-0.009	-0.020	-0.022	-0.022
p75	0.051	0.046	0.070	0.035	0.035	0.050	0.050	0.049
p95	0.239	0.231	0.214	0.144	0.139	0.316	0.293	0.272

[flushleft]*Notes:* In order to put the data on a comparable five-year scale, past import shocks over 1990-1992 are multiplied with the factor 5/2.

Table 9: Counterfactual: Perfect Foresight

	1997-2001	1993-1996	1990-1992
Percentage Change in votes pro-CHN (%)	[-0.055, 0.040]	[-0.227, 0.003]	[-0.031, 0.294]
Always pro-CHN (%)	[55.224, 61.636]	[69.452, 75.549]	[34.674, 40.888]
pro-CHN to against-CHN (%)	[0.007, 1.750]	[0.103, 1.668]	[0.000, 1.423]
Against-CHN to pro-CHN (%)	[0.006, 1.745]	[0.103, 1.520]	[0.000, 1.502]
Always against-CHN (%)	[36.787, 42.944]	[23.080, 28.732]	[57.461, 64.205]

Table 3: MLE Estimates Based on Different Information Sets

	$\hat{a}$ (se.)	$\hat{b}$ (se.)	$\hat{\delta}$ (se.)
		Panel A: 1997-2001	
Minimal information $Z_{it} = \{Share\ Mfg_{it}, \theta_i\}$	0.669 (0.063)	0.254 (0.025)	-0.792 (0.312)
Baseline information $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \theta_i\}$	0.662 (0.062)	0.255 (0.025)	-0.758 (0.277)
Perfect Foresight	0.638 (0.062)	0.254 (0.025)	0.010 (0.210)
		Panel B: 1993-1996	
Minimal information $Z_{it} = \{Share\ Mfg_{it}, \theta_i\}$	0.003 (0.083)	0.648 (0.033)	-0.550 (0.607)
Baseline information $Z_{it} = \{\Delta M_{it} Share\ Mfg_{it}, \theta_i\}$	0.006 (0.082)	0.649 (0.033)	-1.153 (0.544)
Perfect Foresight	-0.002 (0.082)	0.648 (0.033)	-0.474 (0.448)
		Panel C: 1990-1992	
Minimal information $Z_{it} = \{Share\ Mfg_{it}, \theta_i\}$	1.022 (0.102)	-0.224 (0.036)	0.043 (0.350)
Baseline information $Z_{it} = \{\Delta M_{it} Share\ Mfg_{it}, \theta_i\}$	1.005 (0.102)	-0.226 (0.037)	-0.994 (0.286)
Perfect Foresight	1.017 (0.102)	-0.224 (0.037)	-0.659 (0.242)
		Panel D: 1990-2001	
Minimal information $Z_{it} = \{Share\ Mfg_{it}, \theta_i\}$	0.559 (0.044)	0.254 (0.017)	-0.477 (0.220)
Baseline information $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \theta_i\}$	0.561 (0.044)	0.255 (0.017)	-0.857 (0.197)
Perfect Foresight	0.549 (0.043)	0.255 (0.017)	-0.310 (0.146)
		Panel E: 1997-2000	
Minimal information $Z_{it} = \{Share\ Mfg_{it}, \theta_i\}$	0.646 (0.069)	0.255 (0.028)	-0.830 (0.376)
Baseline information $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \theta_i\}$	0.642 (0.068)	0.256 (0.028)	-0.872 (0.332)
Perfect Foresight	0.617 (0.068)	0.255 (0.028)	-0.023 (0.256)

Table 4: Relevance of Manufacturing Employment Share and Past Import Shock in Explaining Future Import Shocks

Sample Period:	1990-2001	1997-2001	1993-1996	1990-1992
Dependent Variable: S_{it+1}	(1)	(2)	(3)	(4)
S_{it}	0.256***	0.735***	0.214***	0.636***
	(0.011)	(0.041)	(0.013)	(0.032)
$Share\ Mfg_{it}$	0.753***	0.683***	0.455***	0.695***
	(0.020)	(0.036)	(0.020)	(0.045)
Observations	5,573	2,572	1,713	1,288
R^2	0.555	0.586	0.673	0.684

[flushleft]*Notes:* In order to put the data on a comparable five-year scale, past import shocks over 1990-1992 are multiplied with the factor 5/2. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Parameter Confidence Sets and Specification Test p-values

Period	CS of a	CS of b	CS of δ	p-value BP	p-value RC	p-value RS	Num obs.
1997-2001	**Panel A:** Minimal information $Z_{it} = \{Share\ Mfg_{it}, \theta_i\}$						
	[0.480, 0.750]	[0.150, 0.245]	[-1.975, -0.137]	0.390	0.390	0.390	2572
	Panel B: Baseline information $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \theta_i\}$						
	[0.465, 0.780]	[0.175, 0.280]	[-1.505, -0.020]	0.350	0.330	0.330	2572
	Panel C: Perfect Foresight $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \Delta M_{it+1} - E[\Delta M_{it+1}\|\Delta M_{it}, Share\ Mfg_{it}], \theta_i\}$						
	–	–	–	0.010	0.010	0.010	2572
1993-1996	**Panel A:** Minimal information $Z_{it} = \{Share\ Mfg_{it}, \theta_i\}$						
	[-0.200, 0.160]	[0.537, 0.672]	[-2.938, 0.887]	0.430	0.430	0.430	1713
	Panel B: Baseline information $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \theta_i\}$						
	[-0.240, 0.180]	[0.560, 0.695]	[-3.500, -0.250]	0.435	0.435	0.435	1713
	Panel C2: Perfect Foresight $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \Delta M_{it+1} - E[\Delta M_{it+1}\|\Delta M_{it}, Share\ Mfg_{it}], \theta_i\}$						
	–	–	–	0.015	0.015	0.015	1713
1990-1992	**Panel A:** Minimal information $Z_{it} = \{Share\ Mfg_{it}, \theta_i\}$						
	[0.800, 1.500]	[-0.375, -0.175]	[-1.200, 1.900]	0.950	0.950	0.950	1288
	Panel B: Baseline information $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \theta_i\}$						
	[0.950, 1.475]	[-0.325, -0.188]	[-1.500, 0.000]	0.300	0.280	0.280	1288
	Panel C: Perfect Foresight $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \Delta M_{it+1} - E[\Delta M_{it+1}\|\Delta M_{it}, Share\ Mfg_{it}], \theta_i\}$						
	[1.000, 1.550]	[-0.300, -0.200]	[-1.550, 0.100]	0.295	0.270	0.270	1288
1990-2001	**Panel A:** Minimal information $Z_{it} = \{Share\ Mfg_{it}, \theta_i\}$						
	[0.495, 0.660]	[0.185, 0.240]	[-1.275, -0.225]	0.220	0.220	0.220	5573
	Panel B: Baseline information $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \theta_i\}$						
	[0.515, 0.673]	[0.210, 0.260]	[-1.500, -0.675]	0.170	0.150	0.150	5573
	Panel C2: Perfect Foresight $Z_{it} = \{\Delta M_{it}, Share\ Mfg_{it}, \Delta M_{it+1} - E[\Delta M_{it+1}\|\Delta M_{it}, Share\ Mfg_{it}], \theta_i\}$						
	–	–	–	0.010	0.010	0.010	5573

Table 6: Parameter Confidence Sets and Specification Test p-values: Heterogeneity by Party

Period	Group	CS of a	CS of b	CS of δ	p-value BP	p-value RC	p-value RS	Num obs.
		(1)	(2)	(3)	(4)	(5)	(6)	
1997-2001	Democrats	[2.400, 4.500]	[0.850, 1.650]	[-4.425, -0.173]	0.810	0.780	0.780	1243
	Republicans	–	–	–	0.010	0.010	0.010	1320
1993-1996	Democrats	[1.000, 3.250]	[1.000, 1.900]	[-6.600, 0.800]	0.910	0.910	0.910	899
	Republicans	–	–	–	0.010	0.010	0.010	810
1990-1992	Democracts	[1.875, 4.125]	[-0.050, 0.750]	[-3.700, -0.400]	0.500	0.500	0.500	785
	Republicans	[-1.200, 1.500]	[-0.350, 0.550]	[-1.450, 1.700]	0.335	0.335	0.335	501
1990-2001	Democracts	[1.500, 3.075]	[0.600, 1.150]	[-3.032, -0.667]	0.790	0.790	0.790	2927
	Republicans	–	–	–	0.010	0.010	0.010	2631

Table 7: Parameter Confidence Sets and Specification Test p-values: Heterogeneity by Tenure

Period	Group	CS of a	CS of b	CS of δ	p-value BP	p-value RC	p-value RS	Num obs.
		(1)	(2)	(3)	(4)	(5)	(6)	
1997-2001	Tenure > median	–	–	–	0.025	0.020	0.020	1377
	Tenure < median	[0.180, 0.740]	[0.220, 0.400]	[-3.237, -0.700]	0.335	0.335	0.335	1195
1993-1996	Tenure > median	[-0.250, 0.250]	[0.475, 0.625]	[-3.625, 0.875]	0.525	0.525	0.525	868
	Tenure < median	[-0.500, 0.250]	[0.650, 0.850]	[-5.600, 0.400]	0.495	0.480	0.480	830
1990-1992	Tenure > median	[0.750, 2.250]	[-0.200, 0.100]	[-4.850, 0.250]	0.885	0.885	0.885	595
	Tenure < median	[0.900, 1.800]	[-0.550, -0.300]	[-1.750, 1.100]	0.495	0.485	0.485	666
1990-2001	Tenure > median	[0.825, 0.825]	[0.200, 0.225]	[-1.635, -0.775]	0.110	0.110	0.110	2533
	Tenure < median	[0.375, 0.375]	[0.250, 0.300]	[-1.312, -0.452]	0.120	0.120	0.120	2982

Table 8: Parameter Confidence Sets and Specification Test p-values: Heterogeneity by Win Margin

Period	Group	CS of a	CS of b	CS of δ	p-value BP	p-value RC	p-value RS	Num obs.
		(1)	(2)	(3)	(4)	(5)	(6)	
1997-2001	Win margin>median	[0.900, 0.900]	[0.125, 0.250]	[-1.255, 0.753]	0.330	0.330	0.330	1284
	Win margin < median	[0.300, 0.300]	[0.225, 0.350]	[-2.693, -0.802]	0.255	0.255	0.255	1288
1993-1996	Win margin>median	[-0.250, 0.250]	[0.475, 0.625]	[-3.175, 1.775]	0.470	0.470	0.470	806
	Win margin < median	[-0.750, 0.250]	[0.600, 0.900]	[-7.600, -0.200]	0.950	0.950	0.950	807
1990-1992	Win margin>median	[0.900, 2.250]	[-0.500, -0.150]	[-3.000, 3.400]	0.875	0.855	0.855	595
	Win margin < median	[0.400, 1.700]	[-0.350, -0.100]	[-4.625, -0.875]	0.590	0.590	0.590	596
1990-2001	Win margin>median	[0.600, 0.825]	[0.125, 0.200]	[-0.820, 0.505]	0.165	0.165	0.165	2588
	Win margin < median	[0.150, 0.375]	[0.250, 0.350]	[-2.388, -0.667]	0.530	0.530	0.530	2589

A Data Appendix

B Monte Carlo Simulation: MLE bias

In this Appendix, we adopt Monte Carlo simulations to illustrate the bias that emerges when estimating ω_t by maximum likelihood in equation (4). We adopt a time horizon of 5 years to define a future import shock, so that $S_{i,t+1}$ corresponds to the increase in import penetration in the next 5 years. The true information set throughout the exercise is what we have defined as Baseline information set $\mathcal{I}_{i,t}^B = \{S_{it}, ShareMfg_{it}\}$. More specifically, $S_{i,t+1}$ is simulated according to the following linear model:

$$S_{i,t+1} = \underbrace{\beta_0 + \beta_1 S_{i,t} + \beta_2 ShareMfg_{it}}_{\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^B]} + \varepsilon_{it},$$

where ε_{it} is the expectational error. To estimate β's, and the residual distribution, we use the data from 1998 and 2001 (XXX clarify for $S_{i,t+1}$). In particular the parameters of the simulated data are set as follows: $\beta_0 = 0$, $\beta_1 = 0.721$, $\beta_2 = 0.184$, $std(\varepsilon_{it}) = 0.525$. We set $a = 0.5$ and $b = 0.3$ and present two sets of results based on $\delta = -1.3$ and $\delta = 0$.

We will evaluate the bias that arises when the econometrician assumes that the politician has Minimal information $\mathcal{I}_{i,t}^M$ and Perfect foresight $\mathcal{I}_{i,t}^P$. For the case of assumed Minimal information set, we estimate ω_t by maximizing the following log-likelihood:

$$\begin{aligned} & ln\mathcal{L}\left(\omega_t | \left\{Y_{i,t}, \theta^i, \mathcal{I}_{i,t}\right\}_{i=1}^N\right) \\ & = \sum_{i=1}^N Y_{i,t} ln\left[\Phi\left(a_t\theta^i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}^M\right]\right)\right] \\ & + (1 - Y_{i,t})\, ln\left[1 - \Phi\left(a_t\theta^i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1}|\mathcal{I}_{i,t}^M\right]\right)\right]. \end{aligned} \tag{A.1}$$

where the expectation is estimated as:

$$S_{i,t+1} = \underbrace{\mu_0 + \mu_1 ShareMfg_{it}}_{\mathbb{E}[S_{i,t+1}|\mathcal{I}_{i,t}^M]} + \upsilon_{it}$$

For the case of perfect foresight the log-likelihood is maximized after replacing the expectation of future shocks with $S_{i,t+1}$. Table C.1 reports the mean and the standard deviation of MLE estimates for 500 simulated datasets. First, the simulations clearly indicate that when the model is

correctly specified, the average parameter estimates are very close to the true parameters. Second, when we assume the politician has perfect foresight, i.e. more information than she actually does, there is a clear attenuation bias. The average $\hat{\delta}$ is -0.813 (the true δ is -1.3). We discussed the intuition for this attenuation bias as related to classical measurement error in the main text. A more nuanced case is the one in which the politician is assumed to have a Minimal information set, i.e. the econometrician assumes that the politician knows strictly less than what she actually knows. In our simulations we find again an attenuation bias, as the average $\hat{\delta}$ is -1.060. While it is not easy to determine, in general, the direction of the bias, it is instructive to discuss the specific result, because it differs from Dickstein and Morales (2018). In our example the true voting decision is taken as follows:

$$\begin{aligned} Y_{it} &= \mathbf{1}\{a_t\theta^i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^B_{i,t}\right] - \xi_{it} \geq 0\} \\ &= \mathbf{1}\{a_t\theta^i + b_t + \delta_t \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^M_{i,t}\right] + \delta_t \left\{\mathbb{E}\left[S_{i,t+1}|\mathcal{I}^B_{i,t}\right] - \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^M_{i,t}\right]\right\} - \xi_{it} \geq 0\} \end{aligned}$$

where $\rho_{it} = \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^B_{i,t}\right] - \mathbb{E}\left[S_{i,t+1}|\mathcal{I}^M_{i,t}\right]$ is the specification error, i.e. the difference between what the econometrician assumes and what the politician actually knows. In Appendix D.5 of Dickstein and Morales (2018) the correlation between the specification error and the assumed expectation, here $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}^M_{i,t}\right]$, is by construction zero, so their estimates display very minor biases. In our case, ρ_{it} and $\mathbb{E}\left[S_{i,t+1}|\mathcal{I}^M_{i,t}\right]$ happen to be negatively correlated (XXXcheck), hence the attenuation bias.

Finally, we consider the special case in which $\delta = 0$. Clearly, if the China shocks truly do not matter for voting decisions, MLE biases are negligible. This is not necessarily the case for case (4) estimates as the bias could go in any direction, but for case (6) attenuation bias will not affect the estimate of a null coefficient.

Table C.1: Maximum Likelihood - Monte Carlo Simulation

	Assumed Information Set	Avg $\hat{a}$ (std.)	Avg $\hat{b}$ (std.)	Avg $\hat{\delta}$ (std.)
$a = 0.5, b = 0.3, \delta = -1.3$	(1) Minimal Information	0.449 (0.066)	0.303 (0.027)	-1.060 (0.047)
	(2) Baseline Information (correct)	0.498 (0.079)	0.319 (0.034)	-1.306 (0.058)
	(3) Perfect Foresight	0.421 (0.090)	0.304 (0.040)	-0.813 (0.190)
$a = 0.5, b = 0.3, \delta = 0$	(4) Minimal Information	0.499 (0.073)	0.300 (0.029)	-0.001 (0.046)
	(5) Baseline Information (correct)	0.499 (0.072)	0.300 (0.029)	-0.000 (0.036)
	(6) Perfect Foresight	0.500 (0.072)	0.300 (0.029)	-0.002 (0.041)